Pre-history of probability

James Franklin

Oxford Handbook of Probability and Philosophy ed. Alan Hájek and Christopher Hitchcock (OUP, 2016), 33-49.

It has been common to begin histories of probability with the calculations of Fermat and Pascal on games of chance in 1654. Those calculations were the first substantial mathematical results on stochastic phenomena, which, according to a frequentist philosophy of probability, are the true subject-matter of the field. But there is great philosophical interest in studying earlier ideas. That is because we see a struggle to understand a range of basic concepts about probability and uncertain evidence, before the straightjacket of a single formalization was imposed. There is still much to learn from seeing how the problems now studied with the aid of probability theory were dealt with in "bare hands" fashion before that formalization was available.

Two concepts: logical/epistemic versus factual/stochastic probability

Certain philosophical distinctions are needed in order to identify what body of ideas in ancient and early modern texts should be regarded as probability. The main distinction needed is that between factual or aleatory probability, on the one hand, and logical or epistemic probability on the other (introductions in Hacking 2006: ch. 2; Mellor 2005: ch. 1; Franklin 2009: ch. 10; finer distinctions in Hájek 2002/2011). *Factual* or *stochastic* or *aleatory* probability deals with chance set-ups such as dice-throwing and coin-tossing, which produce characteristic random or patternless sequences of outcomes. The calculus of probability applies straightforwardly to it.

By contrast, *logical* or *epistemic* probability is concerned with how well-supported a conclusion is by a body of evidence. A concept of logical probability is employed when one says that, on present evidence, the steady-state theory of the universe is less probable than the big bang theory, or that an accused's guilt is "proved beyond reasonable doubt" though not absolutely certain. How probable a hypothesis is, on given evidence, determines the degree of belief it is rational to have in that hypothesis, if that is all the evidence one has that is relevant to it. Views of the nature of such probability range from the objective versions of Keynes

(1921) and Jaynes (2003), holding that the relation of evidence to hypothesis is a matter of strict logic and so a kind of partial implication, to the views of the subjective Bayesian school (e.g. Earman 1992), according to which there is only degrees of belief in propositions, constrained by the laws of probability.

On any of those views, it is unclear how far epistemic probability can be quantified. As an idealization, it is usual to formalise it as a number $P(h|e)$ between 0 and 1, expressing the degree to which evidence *e* supports hypothesis *h*, subject to the usual laws of conditional probability. But it is controversial whether there is such a precise number for arbitrary *e* and *h*. The controversy assumes significance for the prehistory of probability, since most of pre-Pascalian writing that can be considered to be about probability dealt with uncertain evidence, and failure to apply numbers to the evaluation of evidence by early writers is not necessarily a sign of a defect on their part.

The law of evidence and half-proof

Law is a discipline strongly concerned with continuity maintained through a written tradition. Analysis of concepts is central to it. And the complexities involved in real cases keep the concepts grounded in reality and encourage their development. As a result, law was the matrix in which most development of the concepts of probability took place, both those concepts connected with the law of evidence and those involved in aleatory contracts (contracts whose fulfillment depends on chance, like insurance, lotteries, and games of chance).

Ancient Greek law was less constrained than later legal systems by theory and precedent, more a rhetorical free-for-all. Therefore skill in constructing “likely” arguments was at a premium. Aristotle gives an example from Corax’s *Art of Rhetoric* (c. 460BC): “If the accused is not open to the charge, for instance if a weakling is being tried for assault, the defence is that he was not likely to do such a thing. But if he is open to the charge — if he is a strong man — the defence is still that he was not likely to do such a thing, since he could be sure that people would think he *was* likely to do it.” (The word translated as “likely” is *eikos*, literally “like”.) (Aristotle, *Rhetoric*: 1402a17-22; Franklin 2001: 105). Many such arguments are found in the Athenian orators in forensic speeches, some of them based on genuinely likely explanations of the facts rather than shallow appeals to plausibility. Aristotle’s *Rhetoric* constructs a theory of arguments applicable to non-necessary matter, and distinguishes between arguments that happen to persuade and those that ought to persuade. In non-necessary matters, one must use likelihoods (*eikota*) and signs, the likely (*eikos*) being what usually happens. There are also arguments from fallible signs and arguments from examples. Non-deductive arguments that ought to persuade are said to be the subject of the science of dialectic, but Aristotle does not develop that topic.

These ideas were taken over into Latin by writers on rhetoric such as Cicero and Quintilian. Cicero writes "That is probable which for the most part usually happens or which is the general opinion or which has in itself some likeness to these." (Cicero: I.46; Franklin 2001: 116) However the real Roman contribution was not to rhetoric but to systematic law.

Law deals with evidence, hence with the strength and conflict of evidence. The two great ancient systematisations of law, the Jewish *Talmud* and the Roman *Corpus of Civil Law*, both insisted on very high standards of proof for criminal conviction, but recognised that the standard was short of absolute certainty. In Jewish law, conviction requires two witnesses, and the witnesses must give evidence that is not "based on conjecture" but direct. Roman law holds that conviction may not be "on suspicion", for "It is better to permit the crime of a guilty person to go unpunished than to condemn one who is innocent." The charge must be "supported by suitable witnesses, or established by the most open documents, or brought to proof by signs that are indubitable or clearer than light." (Franklin 2001: 5-7) Torture was a feature of Roman law; judges investigating public crimes should not however begin with torture but use whatever "likely and probable arguments" *(argumentis verisimilibus probabilibusque)* were already available. The word "likely" (*verisimilis),* found earlier in Cicero, is used a number of times in the *Corpus of Civil Law* with very much the common meaning it has in modern English, in examples like "it is very likely *(verosimilius)* that the testator had intended rather to point out to his heirs where they could readily obtain forty aurei ... than to have inserted a condition in a trust." (Franklin 2001: 8-9) Both Jewish and Roman law include a concept of "presumption": a proposition taken to be true until the opposite is proved; some but not all of these are likely but uncertain generalisations. (Franklin 2001: 6, 9; Rabinovitch 1973: 112-5)

Ancient Roman law is rather untheoretical, but when the *Corpus* was rediscovered in Italy in the late eleventh century and began to inform Western medieval law, the commentators on it sought to understand the principles and justifications of the Roman rules. In commenting on evidence, they came to realise that there must be pieces of evidence and presumptions of different strengths. The school of Glossators, of the late twelfth century, united several scattered remarks of Roman law under the classification half-proof *(semiplena probatio)*: "It would seem that ... either the plaintiff proves, or not ... I reply that although according to Aristotle it would seem to be an exhaustive division, it is not so according to the laws. There is a medium, namely, half-proof. Say therefore that in such a case proof has been less than full. Then the judge gives the right to complete the case by oath to the plaintiff ... Full proof is by two witnesses, therefore half proof by one." (Azo of Bologna [1966], 254; Franklin 2001: 18) There was never any finer grading attempted, such as quarter-proofs. However, a qualitative grading of the strength of presumptions became standard. In a decree of Pope Innocent III in 1209, the question concerns someone in doubt whether there is some legal impediment to his marriage, and hence wondering if sex with his/her spouse is permitted. It depends on the strength of the evidence:

> It should be distinguished, whether the spouse knows for certain the impediment to the marriage, and then he may not engage in carnal intercourse ... or whether he does not know the impediment for certain, but only believes it ... In the second case, we distinguish, whether his conscience is thus from a light and rash belief, or a probable and discreet one (*ex credulitate levi et temeraria, an probabili et discreta*) … when his conscience presses his mind with a probable and discreet belief, but not an evident and manifest one, he may render the marriage debt, but ought not to demand it. (*Corpus Iuris Canonici*: 2, col. 908; Franklin 2001: 23)

The theory was elaborated by legal writers and applications made. Torture, reintroduced under the influence of Roman law, could be applied only when there was half-proof of guilt (to produce the other half). The smallest quantum of evidence considered was the "indiciolum", a diminutive of "indication" or "sign": If in a murder case a witness testifies that the brother of the accused was an enemy of the deceased, "an indication is certainly not thereby proved, although it makes an *indiciolum* which is very small, and of no strength." (Franklin 2001: 32). The theory of grades of evidence was laid out with many distinctions in the fourteenth-century commentary of Baldus de Ubaldis, which remained standard for centuries. Massive sixteenth-century expansions weighed down by many examples included Mascardus's *On Proofs* (1584) in three volumes and Menochius's *On Presumptions, Conjectures, Signs and Indications* (1587), in two (Franklin 2001: 28-33, 43-6). Although England had a separate legal tradition from the Continent, Latin concepts such as violent presumptions and half-proofs are sometimes found in seventeenth century English law. The extraordinary continuity of legal thought is shown in the fact that nineteenth-century

American texts on evidence law still quote Menochius and Mascardus (Franklin 2001: 60-1, 44).

Law is central to the story of early modern ideas in ways now hard to appreciate. As well as the preservation of ideas within law itself, legal ideas had a much wider currency than today. The originators of mathematical probability were all either professional lawyers (Fermat, Huygens, de Witt) or the sons of lawyers (Cardano and Pascal), and so had some contact with at least the broad concepts of legal thought. Such legal connections were not unusual among intellectual figures. Bacon and Copernicus were also lawyers, Montaigne a judge, Valla a notary, Machiavelli and Arnauld the sons of lawyers, and Petrarch, Rabelais, Luther, Calvin, Donne, Descartes and Leibniz former law students (Franklin 2001: 348-352).

The doctrine of probabilism in moral theology

The Catholic practice of confession led to the development of casuistry, the detailed consideration of "cases of conscience", moral dilemmas on which confessors needed to give advice. The confessional was thought of as something like a miniature court of canon law, and a legal way of thinking was applied to achieving consistency in practical advice (Jonsen and Toulmin 1988). A much-discussed issue in late medieval casuistry was the duty of the conscience in a state of doubt. When one is in doubt about a case, for example when it is unclear whether a rule applies, or when authorities disagree, what should one do? Pope Innocent III laid down an axiom, "In doubtful matters the safer path is to be chosen"; that would be a harsh saying if taken literally, as trivial and over-scrupulous doubts would prevent action.

To distinguish between absolute certainty and the kind of certainty possible in matters of action, the phrase "moral certainty" came to be used. It was introduced by Jean Gerson about 1400, on the basis of a saying of Aristotle that ethics is not a precise science and one must be content with the degree of certainty that the subject admits. One should have, Gerson says, moral certainty that a proposed action is right. To acquire moral certainty, one considers what usually happens, what authorities say, and what one's own learning suggests (Deman 1936: cols 442-443; Franklin 2001: 69-70).

The discussions on the degree of doubt necessary to excuse one from following a doubtful rule are summarized in Sylvester Prierias's standard manual for confessors (about 1515). The entry for "Probable" is:

> "Probable" is used in two ways … Second, as what pertains to opinion. And this in two ways. First, the object of a believed opinion; thus Aristotle (*Topics* I) says that the probable is what seems to be to all, or most, or the wise ... according to the Chancellor [Gerson], what is thus probable is called morally certain ... it is equally vicious for the mathematician to seek the persuasive as for the moralist the demonstrative. (Sylvester Prierias 1515, quoted in Franklin 2001: 71)

Some writers argued for strictness, for example Pope Adrian VI held that a soldier in doubt as to the justice of his king's war ought not to serve in it. But the general tendency was more lax, taking relatively weak doubts to be sufficient to excuse. This tendency reached a culmination in the celebrated doctrine of probabilism, enunciated by the Spanish Dominican Bartolomé de Medina in 1577. According to that doctrine, soon widely accepted in Catholic and to some extent Anglican moral theology, one might follow a course of action that is probable, even if the opposite is more probable. Medina explains what the meaning of the probable is, in terms of having arguments and authorities in favour, in such a way as to allow an opinion to count as probable even when the opposite is more probable:

> Opinions are of two kinds: those are probable which are confirmed by great arguments and the authority of the wise (such as that one may charge interest for delayed payment), but others are completely improbable, which are supported neither by arguments nor the authority of many (such as that one may hold a plurality of benefices) … an opinion is not called probable because there are apparent reasons adduced in its favour, and it has assertors and defenders (for then all errors would be probable opinions), but that opinion is probable which wise men assert, and very good arguments confirm, to follow which is nothing improbable. This is the definition of Aristotle … It could be argued against this that it is indeed in conformity with right reason, but, since the more probable opinion is more in conformity and safer, we are obliged to follow it. Against this is the argument that no-one is obliged to do what is better and more perfect: it is more perfect to be a virgin than a wife, to be a religious than to be rich, but no-one is obliged to adopt the more perfect of those … (De Blic 1930, trans. in Franklin 2001: 74-6; Kantola 1994: 124-130; Maryks 2008: 107-119; Schüssler 2005; Schwartz 2014)

Medina's "probable" is thus close to modern English "arguable". He thinks of "probability" in terms of the reasons (including authorities) in favour of an opinion – in abstraction from the reasons against and the consequent balance of reasons that is normally meant now by the term probability. The question of balance of reasons is further discussed by the most celebrated Spanish Catholic philosopher, Francisco Suárez. He distinguishes usefully between "positive" and "negative" doubt. A doubt is negative when there are no reasons for either side, positive if there are reasons, but they balance. A soldier who knows nothing about the justice of his king's war may trust the king to know that the war is just (assuming the king is of good reputation). "But if the doubt is positive, and there are probable reasons on either side, I believe they are perhaps obliged to inquire into the truth. If they cannot arrive at that, they should follow what is more probable, and aid him who more probably is in the right; for when there is doubt on a matter of fact, and one which concerns the harm of our neighbour or the defence of the innocent, one must follow what appears more probable." (Suárez [1944]: 2: 836; Franklin 2001: 76-9) The distinction is similar to Keynes' concept of the "weight of evidence" for a hypothesis, which may increase as evidence builds up even while its probability, expressing the balance of reasons, stays unchanged (Keynes 1921: ch. 6). Suárez also adds to the conditions for a just war the requirement that there should be "at least a more probable hope of victory, or doubt equally balanced as to defeat or victory, according to the necessity of the state and of the common good." (Suárez [1944]: 2: 822-3)

Probabilism became widely accepted in Catholic moral theology, and further developments included a distinction between the "extrinsic" probability of an opinion (the authorities in its favour) and its "intrinsic" probability (the reasons for it). The morally lax consequences of probabilism were taken to extremes in the work of Juan Caramuel Lobkowitz, who was also the author of one of the earliest books on mathematical probability (Fleming 2006). But probabilism, and casuistry in general, were discredited in popular opinion by the savage attack on them in Pascal's *Provincial Letters* (1656). Pascal caricatured the entire Jesuit order, the main target of his attack, as committed to the thesis that the opinion of a single doctor may render an opinion probable and hence allowable. Pascal's prose style is admirable, but there is little honest argument (Franklin 2001: 94-101; Maryks 2008: 127-134). Pascal's polemical success is the ultimate source of the false belief, still found in Hacking's widely-read *The Emergence of Probability*, that the word "probable" before Pascal meant merely "supported by authorities". That is incorrect because "intrinsic" probability, based on reasons, was always part of the understanding of the probability of opinions.

Evaluating scientific and historical theories

One of the main applications of probabilistic reasoning is to evaluate the strengths of arguments for theories – scientific theories, medical diagnoses, historical claims, and so on. In ancient, medieval and early modern times, the problems with evaluating uncertain theories were widely recognised and the strength of reasons was discussed, although no coherent and accepted overall theory was developed.

Pre-modern and early modern thought had considerable difficulty with the relation of experimental data to scientific theory, with the result that scientific theories were rarely based solidly on observational evidence. One problem was the deductivist model of science according to which the truths of a science should be theorems derivable from axioms evident to reason. That model of science was promoted in Aristotle's *Posterior Analytics*, realized with success in Euclid's geometry, and revived in a new form by rationalists such as Bacon, Descartes and Pascal. Another was that the "experience" that scientific theory was supposed to generalise and explain was thought of as commonly known facts, rather than expensively gathered experimental data; in any case, the institutional funding for expensive data collection was rarely available. Particularly notable by its absence is any attempt to tabulate, summarise and draw inferences from social science or survey data; the first attempt of that kind was Graunt's analysis of London mortality bills in the 1660s (Franklin 2013).

One science did however have a continuous tradition of observation and its relation to theorizing – astronomy. Ptolemy's summary of ancient Greek and Babylonian astronomy includes a notion similar to the averaging of observations to reduce error: to determine the small time by which the length of the year falls short of $365\,^1/_4$ days, he distributes the total observed difference over many years and argues:

> For the error due to the inaccuracy inherent in even carefully performed observations is, to the senses of the observer, small and approximately the same at any [two] observations, whether these are taken at a large or a small interval. However, this same error, when distributed over a smaller number of years, makes the inaccuracy in the yearly motion [comparatively] greater (and [hence increases] the error accumulated over a long period of time), but when distributed over a larger number of years makes the inaccuracy [comparatively] less. (Ptolemy [1984]: 136-7)

Ptolemy also gives some attention to the still-vexed problem of the simplicity of theories. While admitting that in general it is "a good principle to explain the phenomena by the simplest hypotheses possible", he argues that it is different with the heavens because of their divine nature (Ptolemy [1984]: 600-1; Franklin 2001: 137-9). Copernicus too worried incessantly about the observational errors in the ancient, medieval and recent data that he used to support his heliocentric theory, but believed in the end that "so much and so great testimony agree with the mobility of the earth." (Franklin 2001: 146-7)

Both Kepler and Galileo, in arguing for the Copernican hypothesis, speak about the probabilistic relation of evidence to theory with some degree of explicitness. Kepler puts forward the argument that the Copernican theory can do without Ptolemy's complicated system of epicycles, "It is much more probable that there should be some one system of spheres." While geometry is deductive, physics is required when postulating a magnetic-like force moving the planets, so "As is customary in the physical sciences, I mingle the probable and the necessary and draw a plausible *(probabilem)* conclusion from the mixture." Galileo's arguments for Copernicanism include a number of probabilistic ones, including one from the proportionality in the rotations of the spheres:

> … nor do I pretend to draw a necessary proof from this; merely a greater probability (*una maggior probabilità*). The improbability (*l'inverisimile*) is shown for a third time in the relative disruption of the order which we surely see existing among those heavenly bodies whose circulation is not doubtful, but most certain. The order is such that the greater orbits complete their revolutions in longer times, and the lesser in shorter: thus, Saturn … And this very harmonious trend will not be a bit altered if the earth is made to move on itself in twenty-four hours. But if the earth is desired to remain motionless … (Galileo [1967]: 118-9; Franklin 2001: 147-59)

The issues with evidence in biological and social sciences are very different from those in astronomy, because of the variability in the subject matter. That was widely recognised from ancient times. Ancient "sciences" such as divination, astrology and physiognomics particularly needed excuses for their lack of predictive success, and plenty were forthcoming; Ptolemy (and Kepler), for example, defended astrology as only able to predict tendencies, because of the complexity of the planetary influences on fate (Franklin 2001: 165). Medicine was hardly in a better position; as Hippocrates says, "Life is short, art long, opportunity fleeting, experiment deceptive, judgement difficult" (Franklin 2001: 166). The ancient conflict between the Dogmatic and Empiric schools of medicine resembles that between Rationalism and Empiricism, with the Empirics defending experience-based knowledge independent of insight into causes ("Is the peasant, until he has learned from one of the philosophers something of the nature and substance of the soil, and what is the nature and substance of rain and wind, and how they come about, unable to know by experience what seeds to sow at certain times and on what soil, if they are to spring and flourish and attain completion and perfection?" Galen [1965]: 61) The Empirics defended inductive reasoning from what has been observed to happen very many times, and made semi-quantitative distinctions among observed frequencies: "Always, as death in the case of a heart wound; for the most part, as purgation from the use of scammony resin; half the time, as death in the case of a lesion of the dura mater; rarely, as health in the case of a cerebral wound." They also used argument from analogy, with closer similarity justifying more "hope" (Galen [1965]: 31; Franklin 2001: 168-169). There were few improvements in later times. Islamic, European medieval and Renaissance medical writers laid down rules of experimentation, in themselves reasonable, but the variability of the subject matter defeated most attempts to achieve knowledge.

Similar problems arise with the evaluation of historical theories. But in addition, the difficulty with evaluating hypotheses in history is that one cannot experiment and usually cannot collect new evidence. Medieval historians, though often credulous by modern standards, recognised problems with the conflict of authorities, the a priori implausibility of stories, and the authenticity of documents, and asked if there might be rules for establishing the credibility of histories. A specially significant case was the Donation of Constantine, a document forged about 750 but widely believed to be authentic, which purported to be a grant by the Emperor Constantine of the entire Western Roman Empire to the Pope, in perpetuity. Lorenzo Valla in the fifteenth century demolished the case for its authenticity, citing anachronisms but principally using an argument from silence: there is no mention of such a remarkable grant in documents of the time or long after. Melchior Cano, in the sixteenth century, laid down some signs of true histories. He recognised the possibility of tension between historical testimony and the antecedent unlikeliness of a claim, which is the conflict at the basis of Hume's argument in *On Miracles*. Cano says that even serious historians like Pliny might be disbelieved if what they say is too incredible, but one should not be too ready with scepticism as it may be due to one's limited experience. "It would be as if the Mediterranean peoples were to deny the existence of the ocean, or if those who were born on an island in which they saw nothing but hares and foxes should not believe in the lion or the panther, or if, indeed, we should mock at him who speaks of elephants." (Cano [1734]: 89; Franklin 2001: 187-194)

Induction

Contrary to myth, Hume did not discover the problem of induction, though his argument as to why it is unsolvable is original. The problem was one of the many sceptical puzzles posed in ancient times and collected by Sextus Empiricus:

> For, when the Dogmatists attempt to lend credence to a universal by induction from the particulars, in doing this they will consider either all the particulars or only some of them. But if they consider only some, the induction will not be firm, since some of the particulars omitted in the induction may refute the universal; while if they consider all, they will be working at an impossible task, since the particulars are infinite in number and unbounded. (Sextus Empiricus 2: 204 (pp. 160-1); Franklin 2001: 203)

Medieval writers of an Aristotelian tendency understood that something was required to fill the logical gap between particular instances and a universal generalisation, and hoped that some proposition about causality would accomplish the task. Avicenna, for example, said that if we saw the herb scammony being repeatedly followed by purging, we would argue "we have experienced this often and then reasoned that if it were not owing to the nature of scammony but only by chance, this would happen only on certain occasions." Aquinas held that one might make inductive inferences about the future not only about what happens always, but what has a tendency to happen one way rather than the other: "For example, we can conjecture about future effects depending on free choice by considering men's habits and temperaments, which incline them to one course of action." (Aquinas [1952]: 378; Franklin 2001: 203-6) Duns Scotus proposed to justify inductive inference "in virtue of this proposition reposing in the soul: 'Whatever occurs in a great many instances by a cause that is not free, is the natural effect of that cause' ... because a cause that does not act freely cannot in most instances produce an effect that is the very opposite of what it is ordained by its form to produce." (Duns Scotus [1962]: 109-10; Franklin 2001: 206)

The nominalist school, led by William of Ockham, tended to deny such Aristotelian necessary connections, in the light of God's absolute power to produce the opposite of any such alleged necessity. The limit of this tendency was reached in the remarkable work of Nicolas of Autrecourt, the "medieval Hume", soon before the Black Death of 1349. Nicolas argued that no such purported necessities could be reduced to the "necessity of the first principle" – that is, that their denial did not involve a contradiction. A number of Nicolas's propositions were officially condemned by the Church, including one that is a statement of inductive scepticism: "This consequence is not admitted with any evidence deduced from the first principle: 'Fire is brought near to tow and there is no impediment, therefore the tow will burn.'" (Franklin 2001: 210-16)

Certain Jesuit philosophers of the seventeenth century returned to the problem. Given that, as Ockham had said, God could at any moment suspend the laws of nature, why should one presume that he will not do so? The Jesuit cardinal Juan de Lugo argues:

> As long as the contrary be not proved, we always presume that God without miracles or violence allows secondary causes to act naturally, although indeed God could morally do the opposite … Because in case of doubt there suffices a presumption founded on long induction of effects, a bare moral possibility is not sufficient for a prudent doubt or judgement of the opposite. For many things are morally possible, which as long as they are not proved, are not posited, but rather are presumed not to be so … (Juan de Lugo, [1868], 2: 240; Knebel 1992: 189; Franklin 2001: 223)

That still does not exactly explain why God should be on the side of the presumption of regularity. The Jesuit Esparza in the 1650s argued that although God could annihilate the world at any time, the end of the world is *necessarily* a rare event: “it is possible for men to be entirely untouched by any concern for the annihilation of the world, even if God were quite indifferent to that event; because that event belongs to the class of events that can happen but rarely.” (Franklin 2001: 223; Knebel 1992)

Hobbes insisted on the fallibility of induction, but allowed it to be a good bet. He even offered, if only by way of example, numerical odds for it: “though a man have always seen the day and night to follow one another hitherto; yet can he not thence conclude they shall do so, or that they have done so eternally: *experience concludeth nothing universally*. If the signs hit twenty times for one missing, a man may lay a wager of twenty to one on the event; but may not conclude it for a truth.” (Hobbes [1839]: 4:17-18; Franklin 2001: 224)

If that is the correct odds when there are twenty observations for and one against, it is unclear what the odds should be if there are twenty for and none against. The answer to that is no clearer now than in Hobbes’ time.

Two argument forms: “not by chance” arguments and statistical syllogisms

Two argument forms that are crucial to modern quantitative probabilistic inference appear in semi-quantitative form in pre-modern writings. One is the “ruling out of the hypothesis of chance”, which is essential to modern statistical hypothesis testing. If a “null” hypothesis, say that a drug has no effect, would render improbable the observed evidence of many cures following the drug, the hypothesis is said to be ruled out (at some level of significance). Aristotle argues similarly. Do the stars in their daily revolutions move independently or are they all fixed to a sphere? It is observed that those stars that move in large circles (near the celestial equator) take the same time to rotate as those near the Pole Star which rotate in small circles. Aristotle concludes, “If the arrangement was a chance combination, the coincidence in every case of a greater circle with a swifter movement of the star contained in it is too much to believe. In one or two cases it might not inconceivably fall out so, but to imagine it in every case alike is a mere fiction. Besides, chance has no place in that which is natural …” (Aristotle, *On the Heavens:* 289b21-28; Franklin 2001: 133-4; Macdonald 2004). As we saw above, Avicenna proposed to justify induction with such an argument. Similar arguments became a staple of early modern design arguments for the existence of God.

A second argument form is the statistical syllogism (or proportional syllogism or argument from frequencies). If x% of A's are B's, and this is an A, I may conclude with probability x% that this is a B (of course, in the absence of further relevant evidence); if the vast majority of flights that take off land safely, my feeling of safety on takeoff is justified. It is this form of inference that gives rise to the must discussed reference class problem, arising from the fact that an individual case is a member of many classes A, in which the proportion of B's may differ, hence leading to statistical syllogisms with conflicting conclusions. Arguments from what happens "for the most part" were common in ancient rhetoric and elsewhere, particularly explicit examples being found in medicine and in Jewish law (Franklin 2001: 169, 172-5; Gabbay and Koppel 2011). Although there were never any attempts at precise quantitative versions of such arguments, there are some remarkable versions in the work of the fourteenth century mathematical genius Nicole Oresme which involve explicit proportions in infinite sets. In the course of a complex argument against astrology to the effect that the proportion of cases in which it can predict rightly is vanishingly small, Oresme states clearly several times the connection between relative frequency and probability: "It is probable (*verisimile*) that two proposed unknown ratios are incommensurable, because if many unknown ratios are proposed it is most probable that any would be incommensurable with any other"; "if there were some number as to which it were completely unknown what it is or how great it is, and whether it is large or small — as perhaps the number of all the hours that will pass before Antichrist — it will be likely that such a number would not be a cube number. It is similar in games where, if one should inquire whether a hidden number is a cube or not, it is safer to reply that it is not, since that seems more probable and likely (*probabilius et verisimilius*)." (Oresme [1966]: 248-51; Franklin 2001: 140-5; Meusnier 2007b) Unfortunately Oresme's works were little known until recent times.

Aleatory contracts: insurance, options, life annuities, risks

The business world has always had to deal with uncertainty – ships return or not, debtors repay or default, seasons are good or bad – and those risks are reflected in prices. It is not easy to understand how ancient, medieval and modern businesses conceived of risk, since not much was written by businessmen themselves on conceptual matters, and what there is has not been sufficiently studied. However, there is considerable evidence from law and moral theology, which dealt with disputes in business and arguments on the ethics of business dealings.

Ancient Greek and Roman law recognised the maritime loan, a loan advanced for a voyage with repayment only if the ship returned. Variable prices for such loans, for example for differences in seasons, represent an implicit quantification of risk. Roman law says "the price is for the peril." It goes further in accepting that "perils", "hopes" and "expectancies" are entities that can be sold separately and have a price. "Sometimes, indeed, there is held to be a sale even without a thing, as where what is bought is, as it were, a chance (*quasi alea*). This is the case with the purchase of a catch of birds or fish or of largesse showered down. The contract is valid even if nothing results, because it is a purchase of an expectancy (*spei*)." (*Corpus Iuris Civilis, Digest* 18.1.8; Franklin 2001: 259-60) The ancient Jewish law of the Talmud has a similar notion of conditional doubtful claims for which a price may be estimated (Ohrenstein and Gordon 1991). Islamic law, however, forbade any contracts whose fulfilment depended on chance.

In medieval canon law and moral theology, the charging of interest on a loan was condemned as usury. That led to many subtle discussions on what contracts were permissible, for example whether the charging of a fee for late payment might be really a charge for risk instead of genuine interest. It became accepted that prices might reflect common experience of what would probably happen in the future (Piron 2007). Thus Peter Olivi writes, about 1300:

> It is also clear that when someone, by special grace, suffers requisition or sells a crop at a time when it is commonly cheap, but which he firmly intended to keep and sell at a time when it is commonly and probably dearer, he may demand the price which, at the time of requisition, he believed probably would be obtained at the dearer time. (Olivi [1980], 84; Franklin 2001: 266; Ceccarelli 1999)

In discussing the amount of compensation for personal injury, he says "the depriver is required to restore only as much as the probability of profit weighs (*quantum ponderat probabilitas talis lucris*)". In another place, concerning transfer of risks for a fee, he writes "the priceable value of a probability *(appreciabilis valor probabilitatis),* or of a probable hope of profit, from the capital is capable of being traded." Such language clearly envisages the quantification of an expectation, in the sense of a potential outcome weighted by the probability of its happening.

At the same period there was discussion of the licitness and pricing of life annuities. It is argued that such contracts can be fair or not, depending on how the price relates to the probabilities of death: "we see men and women twenty-five years old buying life annuities for a price such that within eight years they will receive their stake back; and although they may live less than those eight years, it is more probable (*probabilius*) that they will live twice that. Thus the buyer has in his favour what happens more frequently and is more probable." (Alexander Lombard [1962], 152-3; Franklin 2001: 270)

Fourteenth century Italy developed insurance contracts, which require the pricing of an expectation. Similar language was used to discuss the price and licitness of insurances (Ceccarelli 2001; Ceccarelli 2007). In the first book on insurance (1488), the Portuguese jurist Santerna explains what the price is for:

> It can be said that the insurer sells only the hope of a future outcome, of which there can well exist a sale . . . from the fact that this hope is uncertain, it might not seem capable of estimation such that in respect of it there could be said to be exceeding of half the just price of its value. But, this is not to be estimated at how much the thing or goods would be worth in case the peril was realised, but at how much the doubtful event should likely (*verisimiliter*) be estimated. (Santerna 1552, 336; Franklin 2001: 277)

The estimation of prices was done by intuitive estimation of all the uncertain factors relevant to the individual case, rather than by collecting statistics. Shakespeare notes the rough quantification of risks by merchants, using the language of odds:

> Or what hath this bold enterprise brought forth
> More than that being which was like to be?
> We all that are engaged to this loss
> Knew that we ventured on such dangerous seas
> That if we wrought out life 'twas ten to one;
> And yet we ventured, for the gain proposed
> Choked the respect of likely peril fear'd. (Henry IV part II, 1.1.178-184)

Shakespeare uses "It's lots to blanks" to mean a near-certainty (Franklin 2001: 280-283; Bellhouse and Franklin 1997). Various kinds of bets, speculations and lotteries were common in business and wider circles, but the contrast between avidity of speculation and lack of theory is remarkable (e.g. Welch 2008).

The moral theologians of the early seventeenth century collected all such contracts under the classification "aleatory contracts", and wrote with clarity on the reality of the probabilistic entities that were being bought and sold. Thus Juan de Lugo (mentioned above in connection with probabilism) wrote, in a chapter on "gaming, wagers and insurance":

> The first condition for the justice of an insurance is, that the price be equal to the peril undertaken; certainly that the price paid for the obligation should be as much as that obligation is worth in the judgement of experts. This price is not a definite amount, but

> has a maximum, mean and minimum, as with buying and selling. As varied circumstances affect the peril, so the just price should be varied. The equality is to be taken from the quantity of the peril at the time of the contract, not after the event. (Juan de Lugo 1652: 2, 447; Franklin 2001: 288)

It is because probability was studied in this moral context that Pascal and Fermat posed questions about dice in a way that is strange to modern ears. The question they ask is about ethics: what is the just division of the stake in an interrupted game of chance? (Coumet 1970; Sylla 2003)

Dice

There is much archeological evidence for ancient and medieval games of chance, but very little theory arose from it. There were occasional comments showing some basic understanding that some outcomes are more frequent than others, such as Aristotle's remark that "To succeed in many things, or many times, is difficult; for instance, to repeat the Koan throw ten thousand times would be impossible, whereas to make it once or twice is comparatively easy," but they were never developed into a quantitative theory.

A remarkable anonymous Italian manuscript of about 1400 poses the problem, very similar to that debated by Pascal and Fermat, of how the stake should be divided in an interrupted game, if one player has won two points and the other none, and the winner is the first to three. It is not an easy problem, as Pascal and Fermat discovered. The author solves it correctly with some complex reasoning from symmetries (and no explicit mention of probabilities or counting of outcomes) (Toti Rigatelli 1985; Franklin 2001: 294-296; background in Meusnier 2007a).

Gerolamo Cardano in the sixteenth century attacked such problems with some generality, and with an understanding of the idea that it involved counting the number of possible outcomes. However his work was often unclear, contained mistakes and was mostly unpublished until the late seventeenth century (Bellhouse 2005). Galileo correctly solved an individual problem involving the throws of three dice, but again his work was not published until centuries later (Franklin 2001: 302).

When Pascal and Fermat came to the problem, some knowledge of probability calculations was already current, but it is hard to discover what it was. The Chevalier de Méré, an amateur mathematician, posed to Pascal this conundrum: if I undertake to throw a 6 in four throws of a die, I have the advantage, with odds in my favour of 671 to 625. But if I undertake to throw a double 6 in 24 throws of two dice, I have the disadvantage. How is that possible? Pascal implies that de Méré was able to work this out for himself, which is not easy. The probabilities were plainly being calculated rather than gained from experience – indeed, no writer of the time suggests any connection between the odds being calculated and long-run relative frequency of throws. No method of working out the probabilities is simple.

Pascal and Fermat apply their well-developed mathematical skills to this problem and to the general problem of determining the just division of the stake in an interrupted game, for any position of the players. They succeed admirably. They say very little, however, about the nature of the entities they are calculating. In their initial letters on the just division of stakes, they merely calculate what would be "impartial" between the players. They appear to have no way of conceptualising a probability except as a just share of a stake, a concept just sufficient for them to deploy the symmetry arguments that result in a numerical solution to the problem. Pascal prefers a mathematical argument involving recursion whereas Fermat prefers enumerating cases, but the difference is purely mathematical. In later letters, Pascal refers to two conditions as "equal and indifferent", while Fermat considers some imaginary cases so as to "make all the chances equal (*rendre tous les hasards égaux*)". That is all that is said. The words "probable" and "probability" do not appear (Franklin 2001: 309-12).

Fermat does not think the topic is very interesting, being merely a source of not especially exciting problems in number theory. Only Pascal realises that something remarkable has been discovered: precisely-calculable necessities in the variable material of chance: "… by thus uniting the demonstrations of mathematics to the uncertainty of chance, and reconciling what seem contraries, it can take its name from both sides, and rightly claim the astonishing title: the Geometry of chance (*aleae Geometria*)." What is calculated, however, is only "what is rightly due" to the players (Pascal, [1964-70]: 2, 1034-5; Franklin 2001: 312).

Philosophical lessons from the pre-history of probability

The story called the "pre-history" of probability has some lessons for the philosophy of probability. They are much the same lessons as arise from a developmental psychology perspective, or a legal perspective, on probability. Reasoning under uncertainty is a very diverse field, comprising everything from subsymbolic risk evaluations in the rat brain to hunches on the demeanour of witnesses to the study of stochastic processes with advanced measure theory. A mathematical model of probability based on counting equiprobable outcomes of dice may or may not be applicable to some or all of the field. That is for debate in individual cases. The "pre-history" of probability, before that model was imposed, is a vast storehouse of examples of the various kinds of non-deductive arguments possible; the arguments are often in pristine form as suggested by real examples, before being deformed by inapplicable theory of one kind or another.

It is particularly significant that the law of evidence, which contributed more than any other discipline to the understanding of uncertain reasoning before Pascal, has almost entirely refused to accept quantification. Apart from some marginal and heavily qualified use of statistics in, for example, DNA identification cases, modern law has strongly resisted any attempts to quantify its probabilistic concepts such as "proof beyond reasonable doubt", and has refused all attempts to apply Bayesian formulas in court (Tribe 1971; Franklin 2011).

So early writers should not be seen as dealing in confused "anticipations" of the later theory of mathematical probability (or hardly ever). They are dealing with matters that in general are as unquantified now as they ever were – the degree to which evidence supports theory, the strength and justification of inductive inferences, the weight of testimony, the combination of pieces of uncertain evidence, the price of risk, the philosophical nature of chance, and the problem of acting in case of doubt (except in those cases where both probability and payoffs are well quantified). Those problems are even now the staple of philosophical discussion of probability.

References

Alexander Lombard [1962], *Le Tractatus de usuris de maître Alexandre d'Alexandria : un traité de morale économique au XIVe siècle,* ed. A.-M. Hamelin (Nauwelaerts, Louvain).

Aquinas, Thomas [1952], *Disputed Questions on Truth,* trans. R.W. Mulligan (Regnery, Chicago)

Aristotle, *On the Heavens.*

Aristotle, *Rhetoric*.

Azo of Bologna [1966], *Lectura super codicem* (Offic. Erasmiana, Turin).

Bellhouse, D. (2005), 'Decoding Cardano's *Liber de Ludo Aleae*', *Historia Mathematica* 32/2: 180-202.

Bellhouse, D. and Franklin, J. (1997), 'The language of chance', *International Statistical Review* 65/1: 73-85.

Cano, M. [1734], *De locis theologicis*, in *Opere* (Typis Seminarii, Padua).

Ceccarelli, G. (1999), 'Le jeu comme contrat et le risicum chez Olivi', in A. Boureau and S. Piron, eds, *Pierre de Jean Olivi (1248-1298): Pensée scolastique, dissidence spirituelle et société* (Vrin, Paris), 239-250.

Ceccarelli, G. (2001), 'Risky business: theological and canonical thought on insurance from the thirteenth to the seventeenth century', *Journal of Medieval and Early Modern Studies* 31/3: 607-658.

Ceccarelli, G. (2007), 'The price for risk-taking: marine insurance and probability calculus in the Late Middle Ages', *Journal Electronique d'Histoire des Probabilités et de Statistique* 3/1.

Cicero, *De Inventione*.

Corpus Iuris Canonici [1979-81], ed. A. Friedberg (Tauchnitz, Leipzig).

Corpus Iuris Civilis, Digest (trans. in A. Watson, *The Digest of Justinian,* University of Pennsylvania Press, Philadelphia, 1998).

Coumet, E. (1970), 'La théorie du hasard est-elle née par hasard?' *Annales: Économies, Sociétés, Civilisations,* 25/3: 574-598.

De Blic, J. (1930), 'Barthélémy de Medina et les origines du probabilisme', *Ephemerides Theologicae Lovanienses* 7: 46-83, 264-291.

Deman, T. (1936), 'Probabilisme', in A. Vacant, E. Mangenot and E. Amann, eds, *Dictionnaire de Théologie Catholique*, XIII, part 1, cols. 417-619 (Letouzey et Ané, Paris).

Duns Scotus, J. [1962], *Opus Oxoniense*, trans. in A. Wolter, *Duns Scotus: Philosophical Writings* (Nelson, Edinburgh).

Earman, J. (1992), *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory* (MIT Press, Cambridge, Mass).

Fleming, J. (2006), *Defending Probabilism: The Moral Theology of Juan Caramuel* (Georgetown University Press, Washington, DC).

Franklin, J. (2001), *The Science of Conjecture: Evidence and Probability Before Pascal* (Johns Hopkins University Press, Baltimore).

Franklin, J. (2009), *What Science Knows: And How It Knows It* (Encounter Books, New York).

Franklin, J. (2011), 'The objective Bayesian conceptualisation of proof and reference class problems', *Sydney Law Review* 33/3: 545-561.

Franklin, J. (2013), 'Probable opinion', in P.R. Anstey, ed, *The Oxford Handbook of British Philosophy in the Seventeenth Century* (Oxford University Press, Oxford).

Gabbay, D.M. and Koppel, M. (2011), 'Uncertainty rules in Talmudic reasoning', *History and Philosophy of Logic* 32/1: 63-69.

Galen [1965], *Three Treatises on the Nature of Science,* ed. M. Frede (Hackett, Indianapolis).

Galileo [1967], *Dialogue Concerning the Two Chief World Systems,* trans. S. Drake (2nd ed, University of California Press, Berkeley).

Hacking, I. (2006), *The Emergence of Probability: A Philosophical Study of Early Ideas about Probability, Induction and Statistical Inference*, 2nd ed (Cambridge University Press, New York).

Hájek, A. (2002, revised 2011), 'Interpretations of probability', *Stanford Encyclopedia of Philosophy*, http://plato.stanford.edu/entries/probability-interpret/

Hobbes, T. [1839], *Human Nature*, in *The English Works of Thomas Hobbes*, ed. W. Molesworth, vol. 4 (John Bohn, London), 1-76.

Jonsen, A.R. and Toulmin, S.E. (1988), *The Abuse of Casuistry: A History of Moral Reasoning* (University of California Press, Berkeley).

Kantola, I. (1994), *Probability and Moral Uncertainty in Late Medieval and Early Modern Times* (Luther-Agricola-Society, Helsinki).

Knebel, S. (1992), ‘Necessitas moralis ad optimum (III): Naturgesetz und Induktionsproblem in der Jesuitenscholastik während des zweiten Drittels des 17. Jahrhunderts,’ *Studia Leibnitiana* 24/2: 182-215.

Keynes, J.M. (1921), *A Treatise on Probability* (Macmillan, London).

Jaynes, E.T. (2003), *Probability Theory: The Logic of Science* (Cambridge University Press, Cambridge).

Juan de Lugo (1652), *De Iustitia et Iure* (Borde, Arnaud, & Rigaud, Lyons).

Juan de Lugo [1868], *Disputationes scholasticae et morales,* ed. J.-B. Fournials (L. Vivès, Paris).

Macdonald, R.R. (2004), Statistical inference and Aristotle’s *Rhetoric*, *British Journal of Mathematical and Statistical Psychology* 57/2: 193–203.

Maryks, R.A. (2008), *Saint Cicero and the Jesuits: The Influence of the Liberal Arts on the Adoption of Moral Probabilism* (Ashgate, Aldershot).

Mellor, D.H. (2005), *Probability: A Philosophical Introduction* (Routledge, London).

Meusnier, N. (2007a), ‘Le problème des partis peut-il être d’origine arabo-musulmane?’ *Journal Electronique d’Histoire des Probabilités et de Statistique* 3/1.

Meusnier, N. (2007b), ‘À propos d’une controverse au sujet de l’interprétation d’un théorème “probabiliste” de Nicole Oresme’, *Journal Electronique d’Histoire des Probabilités et de Statistique* 3/1.

Ohrenstein, R. and Gordon, B. (1991), ‘Risk, uncertainty and expectation in Talmudic literature’, *International Journal of Social Economics* 18/11: 4-15.

Olivi, P.J. [1980], *Un trattato di economia politica francescana : il De emptionibus et venditionibus, de usuris, de restitutionibus di Pietro di Giovanni Olivi* (Istituto storico italiano per il Medio Evo, Rome).

Oresme, N. [1966], *De proportionibus proportionum and Ad pauca respicientes,* ed. and trans. E. Grant (University of Wisconsin Press, Madison, Wisc).

Pascal, B. [1964-70], *Oeuvres Complètes,* ed. J. Mesnard (De Brouwer, Bruges).

Piron, S. (2007), 'Le traitement de l'incertitude commerciale dans la scolastique médiévale,' *Journal Electronique d'Histoire des Probabilités et de Statistique* 3/1.

Ptolemy [1984], *Almagest,* trans. G.J. Toomer (Springer, New York).

Rabinovitch, N.L. (1973), *Probability and Statistical Inference in Ancient and Medieval Jewish Literature* (Toronto University Press, Toronto).

Santerna, P. (Pedro de Santarém) (1552), *De assecurationibus et sponsionibus mercatorum* (Venice, repr. with notes by M. Amzalak, Grémio dos Seguradores, Lisbon, 1971).

Schüssler, R. (2005), 'On the anatomy of probabilism', in J Kraye and R. Saarinen, eds., *Moral Philosophy on the Threshold of Modernity* (Springer, Dordrecht), 91-113.

Schwartz, D. (2014), 'Probabilism reconsidered: deference to experts, types of uncertainty, and medicines', *Journal of the History of Ideas* 75/3, 373-93.

Sextus Empiricus [1996], *Outlines of Pyrrhonism,* in *The Skeptic Way*, trans. B. Mates (Oxford University Press, New York).

Suárez, F. [1944], *Selections from Three Works of Francisco Suárez, S.J.,* trans. G.L. Williams et al (Clarendon, Oxford, 1944).

Sylla, E.D. (2003), 'Business ethics, commercial mathematics, and the origins of mathematical probability', *History of Political Economy* 35, Annual Supplement: 309-337.

Sylvester Prierias, 1515, *Summa summarum,* Bologna.

Toti Rigatelli, L. (1985), 'Il "problema delle parti" in manoscritti del XIV e XV secolo', in M. Folkerts and U. Lindgren, eds, *Mathemata: Festschrift für Helmut Gericke* (Steiner, Stuttgart).

Tribe, E.L. (1971), 'Trial by mathematics: precision and ritual in the legal process', *Harvard Law Review* 84/6: 1329-1393.

Welch, E. (2008), 'Lotteries in early modern Italy', *Past and Present* 199/1: 71-111.